THT2 | 2(B)(d)(ii) $\hat{\mu}_{MLE} \doteq$ 1.45

STAT 206
13 Mar 21

$\hat{SE}(\hat{\mu}_{MLE}) \doteq 0.839$

]] extra
o H

99.9% LCB for μ : $\hat{\mu}_{MLE} - \Phi^{-1}(1-\alpha)\,\hat{SE}(\hat{\mu}_{MLE})$ ①

$\alpha = .001$

$-1.15 = 1.45 - (3.09)\cdot(0.84)$

$\Phi^{-1}(0.999) = \text{qnorm}(0.999) \doteq 3.09$

quantile

dnorm
pnorm
qnorm ← quantile
rnorm

Max lik. conclusion

$\hat{\mu}_{MLE} = 1.45$ is **not** stat. sig. different from 0

99.9% LCB for μ

we can't reject the DA's story

null: $\mu = 0$

-1.15   0   +1.45

↑ percentage points of mortality

Devil's advocate (DA) value for μ

THT 2 | 2(B)(f)(iii) | THT 3 | 2(A)(c)

$$\hat{p}_i = P(\underline{y_i = 1} \mid x_{i1}[5M:LR]\, \underline{\beta})$$

(factor 1 for all mushrooms)

mushroom $i = 1, \ldots, 8124$

$y_i = \begin{cases} 1 & \text{if mushroom } i \text{ is poisonous} \\ 0 & \text{else} \end{cases}$

$\begin{cases} \text{if } y_i = 1 \text{ we want } \hat{p}_i \text{ close to } 1 \\ \text{— } 0 \text{ —————— } 0 \end{cases}$

---

first row of tab.sum(cap.shape):

n = # of mushrooms that have

cap.shape = (1st level of factor) 'b'

alphabetically

$= \hat{p}(\text{poisonous} \mid \text{cap.shape} = \text{'b'})$

mean = E(poisonous | cap.shape = 'b')

outcome, predictor

lm(poisonous ~ cap.shape, ...)

⟷ please fit a linear regression model in which the outcome variable (y) is 'poisonous' & the list of predictor variables includes (in this case) only the factor 'cap.shape' (factor of $\ell = 6$ levels)

$$\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_{i1} + \hat{\beta}_2 x_{i2} + \dots + \hat{\beta}_5 x_{i5} \dots + \cancel{\hat{\beta}_6 x_{i6}}$$

x_{i1}: 1 if cap.shape = 'c', 0

x_{i2}: 1 if c.s = 'f'

x_{i5}: 1 if c.s = 't', 0

R convention

omitted group ~~dummy variable for~~ (level of factor) is the one for the 1st level (alphabetically) of the factor

$$\underset{k+1 \times 1}{\hat{\beta}_{MLE}} = \underset{k+1 \times 1}{\hat{\beta}_{OLS}} = (X^T \underset{n \times k+1}{X})^{-1} X^T y$$

intercept

1 if cap.shape = 'b'

1 if cap.shape = 'c'

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \underset{n \times 6}{X} = \begin{array}{c} \\ \begin{bmatrix} 1 & 1 & 0 & \cdots & 0 \\ 1 & 0 & 0 & \cdots 0 & 1 \\ \vdots & & & & \\ 1 & & & & \end{bmatrix} \end{array} \begin{array}{cc} \text{sum} & \\ \text{'b'} & 1 \\ \text{'x'} & 1 \\ & \vdots \\ & \vdots \end{array}$$

rank$(X) = 6$

1 if cap.shape = 'x'

Sum of last 6 columns equals 1st column

$$Y_i = \beta_0 + \beta_1 x_{i1} + e_i$$

$$\begin{bmatrix} 1 & \\ \vdots & \tilde{x} \\ 1 & \end{bmatrix} \begin{pmatrix} \beta_0 \\ \beta_1 \end{pmatrix}$$

Y

RMSE (root mean squared error) principle

you're going to observe a data vector $y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$

$\left(y_i \in \mathbb{R} \right)$

a realization of the random vector $\underset{\sim}{Y} = \begin{pmatrix} Y_1 \\ \vdots \\ Y_n \end{pmatrix}$

Game: predict y_{n+1}

you get rewarded according to

$$\sqrt{\underline{\underline{E}} \left(\hat{Y}_{n+1} - y_{n+1} \right)^{2}} \, ,$$

" the smaller the better

131 Thm:

the best you can do, having observed $y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$,

is $\hat{Y}_{n+1} = \underset{-}{\bar{Y}}$ (sample mean so far) mean $\bar{y}$

$$= \frac{1}{n} \sum_{i=1}^{n} Y_i$$

using this principle (RAISE ↑)

$$\left(\hat{y} \mid \text{cap.shape} = \text{'b'}\right)_{\text{optimal}} = \bar{y} \text{ when cap.shape} = \text{'b'}$$

(tot. sum) !

$$\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 \begin{cases} 1 & \text{if cap.shape} = \text{'c'} \\ 0 & \text{else} \end{cases}$$

$$+ \cdots + \hat{\beta}_5 \begin{cases} 1 & \text{if } \_\_\_\_ \text{ 'x'} \\ 0 & \_\_\_\_ \end{cases}$$

intercept = $\bar{y}$ of omitted group

$$\left(\hat{y}_i \mid \text{cap.shape} = \text{'b'}\right) = \bar{y} \text{ when cap.shape} = \text{'b'}$$

$$= \hat{\beta}_0$$

$$\left(\hat{y}_i \mid \text{cap.shape} = \text{'c'}\right) = \hat{\beta}_0 + \hat{\beta}_1 = \left(\bar{y} \text{ when c.s} = \text{'c'}\right)$$

$$\hat{\beta}_1 = \left(\bar{y} \text{ when c.s.} = \text{'c'}\right) - \left(\bar{y} \text{ when c.s.} = \text{omitted group}\right)$$

$$\left[\hat{y} \;\middle|\; \overset{\text{factor}}{\cancel{\text{exp.shape}}} = \begin{pmatrix} \text{level} \\ \ell \text{ of} \\ \overset{}{\underset{\text{factor}}{\cancel{\text{exp.shape}}}} \end{pmatrix}\right] = \begin{cases} \bar{y} \text{ when } \overset{\text{factor}}{\cancel{\text{exp.shape}}} = \text{omitted group} \text{ (intercept)} & \ell = 1 \\ \begin{pmatrix} \bar{y} \text{ when} \\ \underset{\text{factor}}{\cancel{\text{e.s}}} = \text{level } \ell \end{pmatrix} - \begin{pmatrix} \bar{y} \text{ when} \\ \overset{\text{factor}}{\cancel{\text{e.s}}} = \\ \text{omitted} \\ \text{group} \end{pmatrix} & \ell > 1 \end{cases}$$

R convention for omitted group when $\ell = 1$